International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-3, September 2019

Multidimensional Space Structure for Adaptable Data Model

Oleksandr Terentyev, Svitlana Tsiutsiura, Tetyana Honcharenko, Tamara Lyashchenko

Abstract*: *The article presents an adaptable data model based on multidimensional space. The main difference between a multidimensional data representation and a table representation used in relational Database Management Systems (DBMSs) is that it is possible to add new elements to sets defining the axes of multidimensional space at any time. This changes the data model. The tabular representation of the relational model does not allow you to change the model itself during the operation of an automated system. Three levels of multidimensional data presentation space are considered. There are axis of multidimensional space, the Cartesian product of the sets of axis values and the values of space points. The five axes of multidimensional space defined in the article (entities, attributes, identifiers, time, modifiers) are basic for the design of an adaptable automated system. It is shown that it is possible to use additional axes for greater granularity of the stored data. The multidimensional space structure defined in the article for an adaptable data model is a flexible set for storing a relational domain model. Two types of operations in multidimensional information space are defined. Relations of the relational model are formed dynamically depending on the conditions imposed on the coordinates of the points. Thus, an adaptable data representation model based on multidimensional space can be used to create flexible dynamic automated information systems.

Keywords*: *multidimensional space, adaptable data model, entity, attribute, relation, identifier, Database Management System, DBMS.

I. INTRODUCTION

The adaptability is one of the most important requirements for information systems for various purposes. This is a characteristic that determines the ability of a system to develop in accordance with the needs of users and business.

Manuscript published on 30 September 2019
***** Correspondence Author
Oleksandr Terentyev, Department of Information Technology Designing and Applied Mathematics, Kyiv National University of Construction and Architecture, Kyiv, Ukraine. Email: terentyev79@ukr.net
Svitlana Tsiutsiura, Department of Information Technology, Kyiv National University of Building and Architecture, Kyiv, Ukraine. Email: svtsutsura@gmail.com
Tetyana Honcharenko*, Department of Information Technology, Kyiv National University of Construction and Architecture, Kyiv, Ukraine. Email: iust511@ukr.net
Tamara Lyashchenko, Department of Information Technology, Kyiv National University of Construction and Architecture, Kyiv, Ukraine. Email: liazschenko@ukr.net

Adaptability is considered quite widely, including in this concept such interconnected non-functional requirements as development ability, flexibility, extensibility, interoperability, etc. [1]. Adaptation of information systems is the process of turning them to changing operating conditions, user needs and business processes both when creating new systems, and when maintaining existing ones. This iterative process can be considered an essential part of the life cycle of an automated system.

The authors of the work [2] describe requirements for the adaptation system. How fast the system can adapt to the new requirements, the implementation of tasks depends on the system.

In the work [3], the author explores the components of an adaptable information system, which consists of two subsystems – a data storage subsystem and an intelligent subsystem. The functioning of the intelligent subsystem depends on stored knowledge and can be modified through a change in stored knowledge. Since the intelligent subsystem must interact with the data storage subsystem, the development of an adaptable data model is an important task determining the ability of a system to develop.

The adaptability of an automated information system is primarily determined by the properties of the data presentation model. Currently, one of the main means of describing and storing data is relational Database Management System (DBMS). This is due to the fact that the relational model is simple in design and implementation compared to other data models (hierarchical, network, object) and has a powerful mathematical apparatus based on set theory, normalization of relationship schemes, and relational algebra.

The articles [4] and [5] are devoted to disadvantages of relational DBMS. The lack of flexibility in the data structure is one of the significant drawbacks that limit the use of relational database to create adaptable systems.

In works [6]–[10] there are traditional structures of relational models in the form of tables that describe various relationship patterns.

The authors of the work [11] describe subject areas that can be modeled using a tabular presentation of the data. It is possible to uniquely fix and define some entities, then for each entity uniquely fix certain attributes and some relationships between different entities in such models.

Articles [12, 13] are devoted the problems of changes to the DBMS when identifying new entities, properties or relationships of the subject area. In this case, the data presentation system must be redesigned, i.e. almost every time a new model of the subject area is created.

Typically, system behavior is considered in a fixed space. The behavior of a system is a change in its states in time, the outcome of which is some result. However, the use of a fixed state space does not allow us to describe the behavior of a complex developing system. Therefore, to create models of adaptable systems that can more adequately reflect the change in the system and its development. It is necessary to use a different way of presenting data, different from the table representation.

II. MAIN RESEARCH

To solve this problem, it is necessary to introduce three equivalent concepts: *entity* (object, thing), *attribute* (property, characteristic) and *relation* (communication, interaction). These three distinguished concepts-categories are enough to represent any information about every subject area [15]. In the process of analyzing a subject area, a certain entity is distinguished and some unique name is assigned to it. If the requirement of uniqueness is not fulfilled, then this will lead to ambiguity and errors. After that, the attributes of the entity are determined, based on which it was selected from the subject area. Each attribute is assigned a specific unique name (name, identifier). Then, all relations of the selected entity with other entities of the given subject area are studied. Each relationship is assigned a unique name (identifier, *ID*).

It is possible to construct a three-dimensional discrete information space <*entity, attribute, relation*> as:

$$<V,\, S,\, O> \tag{1}$$

where V is a set of entities, S is a set of attributes, O is a set of relations.

All three sets V, S, and O for a given subject area are finite sets. The boundaries of such a subject area are large enough and in some cases may increase. The smallest element of this space is a three-dimensional point and is called a mivar, and the space itself is called a mivar space. MIVAR is Multidimensional Informational Variable Adaptive Reality [3]. The coordinates of such mivar are described $M_i\,(v_i,s_i,o_i)$. Mivar is the smallest addressable point of three-dimensional space <*entity, attribute, relation*>.

The formalized description of a three-dimensional data presentation space can be represented as a set:

$$A=\{a_n\},\ n=\overline{1,N} \tag{2}$$

where A is a set of name`s axes of multidimensional space, N is the number of axes of multidimensional space.

The set A contains the structure of a multidimensional space (*MDS*) and collection of the required number of axes:

$$MDS = B_1 \times B_2 \times \ldots \times B_N\,, \tag{3}$$

where MDS is a multidimensional space, B_N is set of axis values a_n, $B_N=\{b_{ni}\},\ n=\overline{1,N},\ i=\overline{1,I_n}$, i_n is element identifier of set of axis values a_n.

The set MDS is the Cartesian product of the sets of axes values of the multidimensional space. The description of the data representation of the subject area occurs using the elements of set B_N.

The data model M of the multidimensional space MDS is described as follows:

$$M=\left\{p_{i_1,i_2,\ldots,i_n}\right\}, \tag{4}$$

where $p_{i_1,i_2,\ldots,i_n}$ is the value of the point of multidimensional space with coordinates $<i_1,i_2,\ldots,i_n>$.

The multidimensional concept of data presentation allows working with dynamic data storage structures, which opens up new possibilities for creating adaptable systems for collecting and processing information.

The multidimensional representation of data allows the use of implicit associative relationships of various concepts and objects. This means that based on the analysis of the structure of the stored data, additional information can be obtained that is not explicitly contained in the database. The concept of a measure of proximity can be introduced in multidimensional space. This distance is either between individual points, or between their sets. It is possible to use a measure of similarity of different structures. This opens up fundamentally new opportunities for CAD or GIS systems using the inference mechanism and laser scanning image processing.

The worldview includes the ideas of an object-oriented approach and reveals new opportunities and prospects. It is possible either to expand the description of the studied objects, or to reduce it. Another feature of the multidimensional information space is that, if necessary, it is possible to enter additional axes and dimensions. Multidimensional space can be divided into separate subspaces and, according to certain rules, can combine separate representations into a single whole.

In terms of relational models, a multidimensional representation is an N-dimensional relational table located in an N-dimensional space in which all the usual relational tables are compiled.

It is necessary to determine the general structure of a multidimensional space to describe an adaptable data model based on a relational model. A relational data model is a set of normalized relations (tables) to which relational algebra operations are applicable [5]. Each relationship includes many attributes and many records, which are determined by the relationship key.

Thus, it is necessary to introduce three axes to describe a relational data model in multidimensional space:

- the axis of entities subject area;
- the axis of entity properties;
- the axis of set of entity record identifiers.

A three-dimensional discrete information space for a relational data model is transformed into the space <*entity, attribute, identifier*>. A formalized description of a such multidimensional data presentation space can be represented as a set:

$$<V,\, S,\, ID> \quad (5)$$

where ID is a set of identifiers.

The space <*entity*, *attribute*, *identifier*> is called the determinant and accordingly the axes forming this space are called determining axes. Then the value of each point in the defining space depends on the parameters of other axes forming a multidimensional space.

The adaptable system changes over time. Therefore, it is necessary to introduce an axis of time that will determine the state of the data model when describing the behavior of an adaptable information system.

The state of the data model at time t is the set of points in a multidimensional space whose values in the determining space with respect to the parameter of the time axis are the last up to a given time t. The state of the data model to the relational model is a set of relations at any given time.

To ensure a multi-user mode of operation and security of stored information, the axis of the system modifiers should be determined, i.e. many users of the system. This axis determines which user has changed state data models.

Thus, the structure of a multidimensional space for an adaptable data model consists of five main axes:

1. The set of entities subject area is described as follows:

$$V = \{v_i\},\ i = \overline{1, N_V} \quad (6)$$

where V is a set of objects, N_V is the number of entities.

2. The set of attributes of entities subject area is described as follows:

$$S = \{s_i\},\ i = \overline{1, N_S} \quad (7)$$

where S is a set of attributes of entities subject area, N_S is the number of attributes of entities.

3. The set of entity record identifiers is described as follows:

$$ID = \{id_i\},\ i = \overline{1, N_{id}} \quad (8)$$

where ID is a set of entity record identifiers, N_{id} is the number of identifiers. An identifier extracts a specific record from a relationship in each entity. The record identifier is not repeated in every respect and is unique.

4. The set of times the state changes of the data model is described as follows:

$$T = \{t_i\},\ i = \overline{1, N_t},\quad t_i < t_{i+1} \quad (9)$$

where T is a set of times, inequality $t_i < t_{i+1}$ means that event t_i did not occur before event t_{i+1}.

5. The set of modifiers is described as follows:

$$U = \{u_i\},\ i = \overline{1, N_U} \quad (10)$$

where U is set of modifiers or identifiers of the users, N_U is number of modifiers, u_i is identifier of the user who made the change in the value of the point of the multidimensional space.

The structure of a multidimensional space taking into account formulas (6)-(10) will have the following form:

$$MDS = V \times S \times ID \times T \times U \quad (11)$$

Accordingly, any point of a multidimensional space $p \in MDS$ will have the following coordinates:

$$p = <v, s, id, t, u> \quad (12)$$

Fig. 1 shows the subspace $<V,\, S,\, ID>$ that defines the relational data model, and the highlighted point stores the value of the attribute v_i with identifier id_j of the relation s_k .

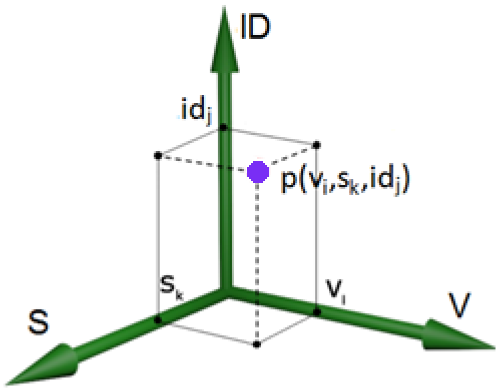

Fig. 1.The data presentation information subspace

The multidimensional space for an adaptable data model stores relationships of a relational model of the subject area. The scheme of operations in multidimensional information space is presented in fig. 2. Therefore, to obtain the query result, it is advisable to divide the operations into two types.

The first type of operations is operations on elements of axes of multidimensional space. There are operations with coordinates of points. These operations are performed on the coordinates of points in a multidimensional space without analyzing the values of these points. These are operations on allocation of various subspaces corresponding certain conditions along the axes (operations to obtain slices of multidimensional space).

The second type of operations is relational algebra operations (SQL).

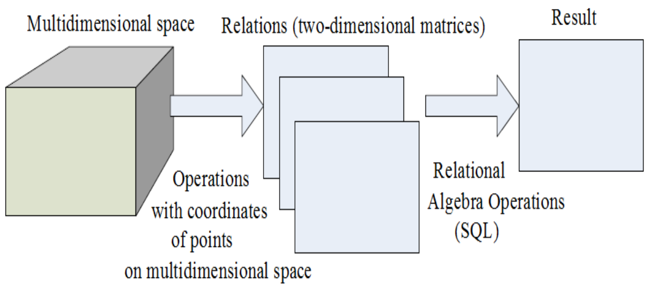

Fig. 2.The scheme of operations in multidimensional information space

Therefore, the function of determining the value of a point of a multidimensional space can be represented:

$$m_p = M(p) \quad (13)$$

Operations with the coordinates of points make it possible to select the required subspace for solving the problem of the subject area. Then, the selected subspace is transformed into the relations of the subject domain by repeatedly applying the function $M(p)$ to all points of the selected subspace. Next, relational algebra operations can be applied to the obtained relations to determine the query result.

The multidimensional information space *MDS* can be represented in matrix form as follows:

$$M = \left\| m_p \right\| \tag{14}$$

The multidimensional space taking into account formulas (12)-(14) will have the following form:

$$M = \left\| m_{v,s,id,t,u} \right\| \tag{15}$$

The matrix consists of the values of the points of multidimensional space. If the point is not included in the description of the subject area, then its value is “null”.

The algebra of multidimensional matrices contains 6 operations: access area, union, intersection, difference of subspaces of multidimensional matrices, section and slice. It is necessary to describe each of the operations for multidimensional matrices.

The operation "Access area" for the user is called a multidimensional matrix of the following form:

$$M_{u_i} = \Phi\left(M_{MDS}\right) = \left\| m_{v,s,id,t,u} \right\|, \quad \phi\left(v,s,id,t,u\right) = true, \tag{16}$$

where $\phi\left(v,s,id,t,u\right)$ is the function of checking the ownership of the matrix element with the coordinates $p = < v,s,id,t,u >$ of the access area defined by the user. The access area makes available that part of the multidimensional space that the user has access.

Let two multidimensional matrices M_A and M_B are given as follows:

$$M_A = \left\| m_{v_A,s_A,id_A,t_A,u_A} \right\|, \quad v_A \in V_A \subseteq V,\ s_A \in S_A \subseteq S,\ id_A \in ID_A \subseteq ID, \quad t_A \in T_A \subseteq T,\ u_A \in U_A \subseteq U, \tag{17}$$

$$M_B = \left\| m_{v_B,s_B,id_B,t_B,u_B} \right\|, \quad v_B \in V_B \subseteq V, s_B \in S_B \subseteq S,\ id_B \in ID_B \subseteq ID, \quad t_B \in T_B \subseteq T,\ u_B \in U_B \subseteq U, \tag{18}$$

The operation "Union" of subspaces of multidimensional matrices is defining as follows:

$$M_D = M_A \cup M_B = \left\| c_{v_D,s_D,id_D,t_D,u_D} \right\|, \quad v_D \in V_A \cup V_B, s_D \in S_A \cup S_B, id_D \in ID_A \cup ID_B, \quad t_D \in T_A \cup T_B,\ u_D \in U_A \cup U_B. \tag{19}$$

The operation "Intersection" of subspaces of multidimensional matrices is defined as follows:

$$M_D = M_A \cap M_B = \left\| c_{v_D,s_D,id_D,t_D,u_D} \right\|, \quad v_D \in V_A \cap V_B, s_D \in S_A \cap S_B, id_D \in ID_A \cap ID_B, \quad t_D \in T_A \cap T_B,\ u_D \in U_A \cap U_B. \tag{20}$$

The operation "Difference" of subspaces of multidimensional matrices is defining as follows:

$$M_D = M_A \setminus M_B = \left\| c_{v_D,s_D,id_D,t_D,u_D} \right\|, \quad v_D \in V_A \setminus V_B, s_D \in S_A \setminus S_B, id_D \in ID_A \setminus ID_B, \quad t_D \in T_A \setminus T_B,\ u_D \in U_A \setminus U_B. \tag{21}$$

Since all multidimensional matrices are parts of the same multidimensional space, then in different matrices the elements having the same indices have the same values. Therefore, when operations combining, intersecting, and difference between subspaces of multidimensional matrices, there is no need to check the values of points in space.

The result of these operations will always be a multidimensional matrix that describes the corresponding certain part of the general multidimensional space.

The set of matrix elements with a fixed value of the indices is called the section with orientation.

The operation "Section" of a multidimensional matrix is a multidimensional matrix of the following form:

$$M_\theta = \theta_f \left\| m_{i_1,\dots,i_n} \right\| \tag{22}$$

where *f* is function for checking the logical condition for the indices of the original multidimensional matrix $i_1,\dots,i_n,\ n \le N$.

The operation "Section" exists at various levels. The first-level section is an n-dimensional matrix in which the value of one index is fixed. The second level section is called an N-dimensional matrix, in which the values of two indices are fixed, etc.

The operation "Slice" of a multidimensional matrix is a multidimensional matrix of the following form:

$$M_\psi = \psi_f \left\| m_{i_1,\dots,i_n} \right\| \tag{23}$$

where $f(i_1,\dots,i_n)$ is function for checking a logical condition that does not include equality of indices to a certain value.

III. RESULT

The operations of the algebra of multidimensional matrices can be considered as an extension of the operations of relational algebra and the existing language SQL. These operations are used to dynamically form the relations of the relational model stored in a multidimensional space, depending on the conditions imposed on the coordinates of the points. Fig. 3 shows the multidimensional space in the form of a three-dimensional cubic matrix.

Retrieval Number: C6318098319/2019©BEIESP
DOI:10.35940/ijrte.C6318.098319
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

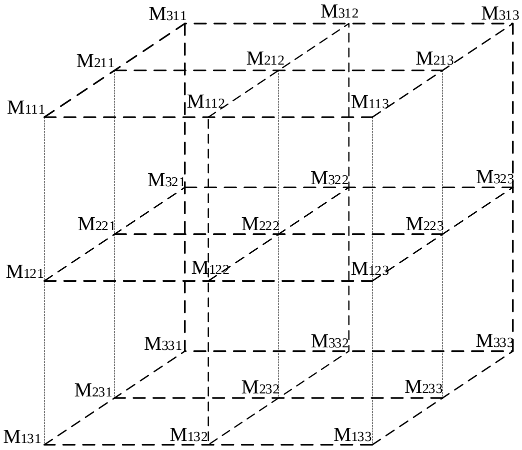

Fig. 3.The graphical representation of the multidimensional space in the form of a three-dimensional cubic matrix

The graphical representation of operation "Section" is shown in fig. 4. The example shows the result of the operation "Section", when the third coordinate is fixed in the three-dimensional matrix. The operation result is a two-dimensional matrix.

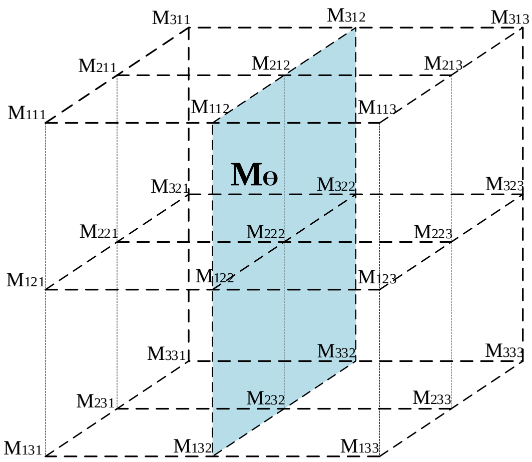

Fig.4.Graphical representation of operation "Section"

The graphical representation of operation "Slice" is shown in fig. 5. The example shows the result of the operation "Slice".

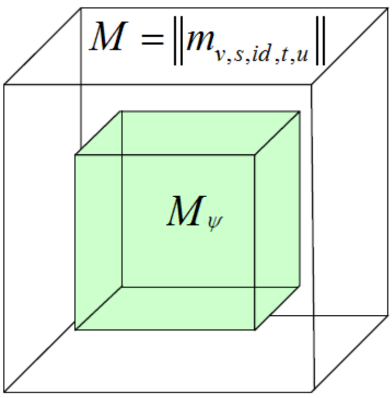

Fig.5.Graphical representation of operation "Slice"

The operations of the algebra of multidimensional matrices are intended for processing indices of matrix elements. These operations do not analyze the values of the matrix elements, but form the relationships of the relational model depending on the conditions imposed on the indices. Algebra is defined on a set of multidimensional matrices obtained from a multidimensional space representing the entire relational data model. This ensures that when applying the introduced operations on multidimensional matrices, we will not go beyond the multidimensional space representing the entire relational data model.

IV. CONCLUSION

In the work, the multidimensional space structure for adaptable data model and introduced multidimensional matrix algebra to work with this space.

As a result of the study, the main five axes of multidimensional space (entities, attributes, identifiers, time, and modifiers) for the design of an adaptable automated system are identified.

The operations of the algebra of multidimensional matrices can be considered as an extension of the operations of relational algebra and the existing language SQL. These operations are used to dynamically form the relations of the relational model stored in a multidimensional space, depending on the conditions imposed on the coordinates of the points. Then, to obtain the final result of the query on the obtained relations, one must apply the operations of relational algebra.

The results of the study have shown that an adaptable data representation model based on multidimensional space can be used to create flexible dynamic automated information systems.

REFERENCES

1. M.A. Serrano, C. Calero, M. Piattini, Experimental validation of multidimensional data models metrics, in*: 36th Annual Hawaii International Conference on System Sciences, 2003, IEEE*, 2003, pp. 1–7.
2. S.S. Shadrin, O.O. Varlamov, A.M. Ivanov, "Experimental Autonomous Road Vehicle with Logical Artificial Intelligence", in *Journal of Advanced Transportation,* 2017: 1–10. doi:10.1155/2017/2492765. ISSN 0197-6729
3. O. O. Varlamov, "MIVAR: Transition from Productions to Bipartite Graphs MIVAR Nets and Practical Realization of Automated Constructor of Algorithms Handling More than Three Million Production Rules". Cornell University, 2011, arXiv:1111.1321. Bibcode:2011arXiv1111.1321V
4. D.A. Chuvikov, N.A. Kazakova, O.O.Varlamov, A.V. Goloviznin, "3D Modeling and 3D Objects Creation Technology Analysis for Various Intelligent Systems" in *International Journal of Advanced Studies*. 4(4): 16, 2015 doi:10.12731/2227-930x-2014-4-3.
5. Milani, M. Multidimensional Ontologies for Contextual .ality Data Speci€cation and Extraction. PhD in Computer Science .esis, Carleton University, 2017. http://people.scs.carleton.ca/_bertossi/papers/mostafaFinal.pdf
6. Morak, M. .e Impact of Disjunction on Reasoning under Existential Rules. PhD in Computer Science .esis, University of Oxford, 2015.
7. Jensen, Ch. S., Bach Pedersen, T. and .omsen, Ch. Multidimensional Databases and Data Warehousing. Morgan &Claypool, 2010.
8. Malaki, A., Bertossi, L. and Rizzolo, F. Multidimensional Contexts for Data .ality Assessment. In *Proc. of the Alberto Mendelzon International Workshop on Foundations of Data Management (AMW)*, 2012. CEUR-WS, Vol-866.

9. Milani, M. and Bertossi, L. Ontology-Based Multidimensional Contexts with Applications to .ality Data Speci€cation and Extraction. In Proc. of the International Symposium on Rules and Rule Markup Languages for the Semantic Web (RuleML), Springer LNCS 9202, 2015, pp. 277-293.
10. Yaghmaie, M., Bertossi, L. and Ariyan, S. Repair-Oriented Relational Schemas for Multidimensional Databases. In *Proc. of the International Conference on Extending Database Technology (EDBT)*, 2012.
11. Franconi, E. and Sa.ler, I. A DataWarehouse Conceptual Data Model for Multidimensional Aggregation. *In Proc. Of the International Workshop on Design and Management of Data Warehouses (DMDW)*, 1999, Article No. 13.
12. S. Nagpal, A. Gosain, S. Sabharwal, Complexity metric for multidimensional models for data warehouse, *in: CUBE International Information Technology Conference, ACM*, 2012, pp. 360–365.
13. W.-K. Chen, *Linear Networks and Systems* (Book style). Belmont, CA: Wadsworth, 1993, pp. 123–135. M. Goeken, R. Knackstedt, Multidimensional reference models for data warehouse development, in: *International Conference on Enterprise Information Systems, Citeseer*, 2007, pp. 347–354.
14. N. Stefanovic et al, "Object-Based Selective Materialization for Efficient Implementation of Spatial Data Cubes," *IEEE Transactions on Knowleadge and Data Engineering, vol. 12*, n. 6, p. 938-958, 2000.
15. M. Body et al, "Handling Evolutions in Multidimensional Structures" *IEEE 19th Int. Conf. on Data Engi-neering (ICDE)*, March 5-8, Bangalore, India, 2003.
16. P.A. Longley, M.F. Goodchild, D.J. Magurie, D.W. Rhind., *Geographical Information Systems,Principals and Technical Issues*, 2nd ed. Wiley & Sons, 2002.
17. P. Rigaux, M. School and A.Voisard, *Spatial Databases with Application to GIS*, Morgan Kaufman 2002.
18. V. Mihaylenko, T. Honcharenko, K. Chupryna, Yu. Andrashko, S. Budnik Modeling of Spatial Data on the Construction Site Based on Multidimensional Information Objects the in '*International Journal of Engineering and Advanced Technology (IJEAT)*' at Volume-8 Issue-6, August 2019, available at http://www.ijeat.org/download/volume-8-issue-6/.

AUTHORS PROFILES

Oleksandr Terentyev, Kyiv National University of Construction and Architecture, Dr. Sc., Professor, Department of Information Technology Design and Applied Mathematics, Deputy Dean of the Faculty of Automation and Information Technologies of KNUBA, Chief Scientific Advisor, Senior Researcher at the State Enterprise "Research Institute of Construction Production named by. V.S. Balytskyi Minregion of Ukraine, Laureate of the Academy of Construction of Ukraine academician M.S. Budnikova. Author 4 monographs, 2 textbooks and more than 150 scientific researches and papers.

Svitlana Tsiutsiura, Kyiv National University of Construction and Architecture, Dr. Sc., Professor, Head of Department of Information Technology, Faculty of Automation Information Technology, Academician of the Academy of Construction of Ukraine. Author over 250 works including 6 manuals, 2 monographs. Scientific interests: management of innovation projects modernization of energy-intensive industries; methodology of budget programs; metrology, measurement bases, standardization and certification.

Tetyana Honcharenko, Kyiv National University of Construction and Architecture, PhD, Associate Professor, Department of Information Technology, author of 25 scientific publications, jury`s member of the All-Ukrainian Student Olympiad on programming. Scientific interests: Computer Aided Design, Building Information Modeling, Geographical Information Systems, multidimensional information objects, construction site master plans, C++, C# app development.

Tamara Lyashchenko, Kyiv National University of Construction and Architecture, Senior Lecturer, Department of Information Technology. Author over 50 scientific works, technical editor of the Ukrainian scientific journal "Management of the development of complex systems". Scientific interests: multidimensional information objects, Informatization of higher education..